

Journal of Medical Imaging

MedicalImaging.SPIEDigitalLibrary.org

Magnetic resonance imaging-based pseudo computed tomography using anatomic signature and joint dictionary learning

Yang Lei
Hui-Kuo Shu
Sibo Tian
Jiwoong Jason Jeong
Tian Liu
Hyunsuk Shim
Hui Mao
Tonghe Wang
Ashesh B. Jani
Walter J. Curran
Xiaofeng Yang

SPIE.

Yang Lei, Hui-Kuo Shu, Sibotian, Jiwoong Jason Jeong, Tian Liu, Hyunsuk Shim, Hui Mao, Tonghe Wang, Ashesh B. Jani, Walter J. Curran, Xiaofeng Yang, "Magnetic resonance imaging-based pseudo computed tomography using anatomic signature and joint dictionary learning," *J. Med. Imag.* 5(3), 034001 (2018), doi: 10.1117/1.JMI.5.3.034001.

Magnetic resonance imaging-based pseudo computed tomography using anatomic signature and joint dictionary learning

Yang Lei,^a Hui-Kuo Shu,^a Sibo Tian,^a Jiwoong Jason Jeong,^a Tian Liu,^a Hyunsuk Shim,^{a,b} Hui Mao,^b Tonghe Wang,^a Ashesh B. Jani,^a Walter J. Curran,^a and Xiaofeng Yang^{a,*}

^aEmory University, Winship Cancer Institute, Department of Radiation Oncology, Atlanta, Georgia, United States

^bEmory University, Winship Cancer Institute, Department of Radiology and Imaging Sciences, Atlanta, Georgia, United States

Abstract. Magnetic resonance imaging (MRI) provides a number of advantages over computed tomography (CT) for radiation therapy treatment planning; however, MRI lacks the key electron density information necessary for accurate dose calculation. We propose a dictionary-learning-based method to derive electron density information from MRIs. Specifically, we first partition a given MR image into a set of patches, for which we used a joint dictionary learning method to directly predict a CT patch as a structured output. Then a feature selection method is used to ensure prediction robustness. Finally, we combine all the predicted CT patches to obtain the final prediction for the given MR image. This prediction technique was validated for a clinical application using 14 patients with brain MR and CT images. The peak signal-to-noise ratio (PSNR), mean absolute error (MAE), normalized cross-correlation (NCC) indices and similarity index (SI) for air, soft-tissue and bone region were used to quantify the prediction accuracy. The mean \pm std of PSNR, MAE, and NCC were: 22.4 ± 1.9 dB, 82.6 ± 26.1 HU, and 0.91 ± 0.03 for the 14 patients. The SIs for air, soft-tissue, and bone regions are 0.98 ± 0.01 , 0.88 ± 0.03 , and 0.69 ± 0.08 . These indices demonstrate the CT prediction accuracy of the proposed learning-based method. This CT image prediction technique could be used as a tool for MRI-based radiation treatment planning, or for PET attenuation correction in a PET/MRI scanner. © 2018 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: [10.1117/1.JMI.5.3.034001](https://doi.org/10.1117/1.JMI.5.3.034001)]

Keywords: MRI-based treatment planning; pseudo computed tomography; dictionary learning; feature selection.

Paper 18093R received May 12, 2018; accepted for publication Aug. 6, 2018; published online Aug. 24, 2018.

1 Introduction

Magnetic resonance imaging (MRI) has several important advantages over computed tomography (CT) for radiation treatment planning.¹ Chiefly, MRI significantly improves soft-tissue contrast over CT, which increases the accuracy and reliability of target delineation in multiple anatomic sites. A potential treatment planning process incorporating MRI as the sole imaging modality could eliminate systematic MRI-CT coregistration errors, reduce cost, minimize radiation exposure to the patient, and simplify clinical workflow.² Despite these advantages, MRI contains neither unique nor quantitative information on electron density, which is needed for accurate dose calculations and generating reference images for patient setup. Therefore, in this proposed MRI-only-based treatment planning process, electron density information that is otherwise obtained from CT must thus be derived from the MRI by synthesizing a so-called pseudo CT (PCT).^{3,4}

Accurate methods for estimating PCT from MRI are crucially needed. The existing estimation methods can be broadly classified into the following four categories:

Atlas-based methods: These methods use single or multiple atlases with deformable registrations.^{5–8} They can also incorporate pattern recognition techniques to estimate electron density.^{9,10} The main drawback is that

their accuracy depends on that of the intersubject registration.

Classification-based methods: These methods first manually or automatically segment MR images into bone, air, fat and soft-tissue classes, and then assign a uniform electron density to each class.^{11–16} However, these methods may not reliably be able to differentiate bone from air due to the ambiguous intensity relationship between bone and air using MRI.

Sequence-based methods: These methods create PCTs by using intensity information from standard MR, specialized MR sequences such as the ultrashort echo time (UTE), or a combination of the two.^{17–25} However, the current image quality of UTE sequences is unsatisfactory for accurate delineation of blood vessels from bone; both similarly appear dark.^{14,25}

Learning-based methods: In these methods, a map between CT and MRI is learned by a training data set and then used for predicting PCT for the target MRI.^{4,26–32} Recent work have proposed a Gaussian mixture regression (GMR) method for learning from UTE MR sequences.^{19,33–36} Due to lack of spatial information, the quality of GMR method was suboptimal. Li et al.³⁷ proposed convolutional neural networks. Random forests have also been used for

*Address all correspondence to: Xiaofeng Yang, E-mail: xyang43@emory.edu

learning the map.^{38–41} Huynh et al.⁴¹ used the structural random forest to estimate a PCT image from its corresponding MRI. However, the redundant features often contain several biases, which may affect the training of decision tree and thus cause prediction errors. Recently, another group of learning-based methods are the dictionary-learning-based methods.^{3,42,43} Andreassen et al.^{3,42} used a patch-based method to generate PCTs based on conventional T1-weighted MRI sequences. They refined this method for use in MR-based pelvic radiotherapy. Due to the parallelizable iteration, a fast patch-based PCT synthesis accelerated by graphics processing unit (GPU) for PET/MR attenuation correction was proposed.⁴³ Aouadi et al.⁴⁴ proposed a sparse patch-based method applied to MR-only radiotherapy planning.

Since recent dictionary-learning-based methods have not taken into consideration patient-specific features when representing an image patch, the purpose of this work is to address this issue by the following improvement.

1. Position and deformation balance: Three-dimensional (3-D) rotation invariant local binary pattern (3-D RILBP) features are innovatively crafted with multiradii level sensitivity and multiderivate image modeling to balance rigidity of image deformation and sensitivity of small position difference.
2. Anatomic signature: A feature selection method is introduced to identify the informative and salient features that are discriminative for clustering bone, air, and soft tissue of each voxel, by minimizing the least absolute shrinkage and selection operator (LASSO) energy function. The identified features, known as the anatomical signature, are used to perform dictionary learning.
3. Joint dictionary learning: To address the challenge of training large amounts of data, for each patch of new arrival MRI, the local self-similarity is applied to restrict the training domain. Then, a joint dictionary learning model is proposed to sparsely represent the new MRI patch by simultaneously training the dictionary of MRI anatomic signature and CT intensity. The alternative direction method (ADM)⁴⁵ is introduced to train the coupled dictionaries of MR and CT by feeding both to distributed optimization. The basic pursuit (BP)⁴⁶ method combined with well-trained dictionaries is utilized for the sparse representation of a new MRI patch.

This paper is organized as follows: we provide an overview of the proposed MRI-based PCT framework in the methods, followed by the details on anatomic signature, construction of the coupled similar patch-based dictionary via joint dictionary learning, and the prediction of target CT by sparse representation reconstruction. We compare state-of-art methods based on dictionary learning with our approach in the results and conclude that our dictionary-learning-based PCT estimation framework could be a useful tool for MRI-based radiation treatment planning for rigid structures.

2 Methods

2.1 System Overview

For a given pair of MR and CT training images, the CT image was used as the regression target of the MRI image. We assumed that the training data have been preprocessed by removing noise and uninformative regions and have been aligned and normalized using mean and standard deviation. In the training stage, for each 3-D patch of MR image, we extracted 3-D RILBP features with multiradii-level sensitivity and multiderivate image modeling including the original MR image and the derivate image filtered through standard deviation and mean filters. The mean and standard filtered images were obtained by assigning each voxel a value corresponding to the mean and standard value of the voxels in a spherical neighborhood around the voxel, respectively. The CT label of each patch's central voxel was obtained using fuzzy C-means labeling. The most informative features corresponding to cluster the labels were identified by an LASSO operator; their statistical discriminative power was evaluated by each one's Fisher's score.⁴⁷ For each patch of new arrival MRI, the training data were collected by searching similar neighbors within a bounded region. Then, we used a joint dictionary learning method to simultaneously train the sparse representation dictionary of both anatomic signature and CT intensities within the searched region. Last, the sparse coefficients of this patch were generated by BP⁴⁶ methods under the dictionary of MRI features. Finally, we used these coefficients to reconstruct the PCT image under the well-trained dictionary of CT. A brief workflow of our method is shown in Fig. 1.

2.2 Preprocessing

A rigid intrasubject registration was used to align each pair of brain CT and MR images of the same subject. All pairs were aligned onto a common space by applying a rigid-body intersubject registration. All the training pairs were further registered to a new test MR image. We performed denoising, inhomogeneity correction, and histogram matching for MR images. To reduce the computational cost of training the models, a mask was performed by removing the most useless regions, i.e., the air region outside the body contour, in all MRI images.

2.3 Feature Extraction

Based on the approach proposed by Pauly et al.,⁴⁸ in which two-dimensional (2-D) LBP-like features were used for predicting organ bounding boxes in Dixon MR image sets, a 3-D rotation invariant version was implemented for this task. The 3-D version differs from 2-D version. The neighborhood voxels are not defined by a circular neighborhood but instead by a spherical one, which can be solved by the RILBP operator proposed by Fehr and Burkhardt.⁴⁹ To address MRI artifacts, we used multiderivate image modeling to extract features with different image models, and also used multiple spatial scales with fixed patch size to achieve the feature's multiradius level sensitivity as recommended previously.⁴¹ For 3-D RILBP feature extraction, we used a cube with size of $3 \times 3 \times 3 \text{ mm}^3$ and a sphere with a 3-mm radius for the first scale and then added 2 and 4 mm to those values for the second and third scales, respectively. The 3-D RILBP feature maps of multilevel

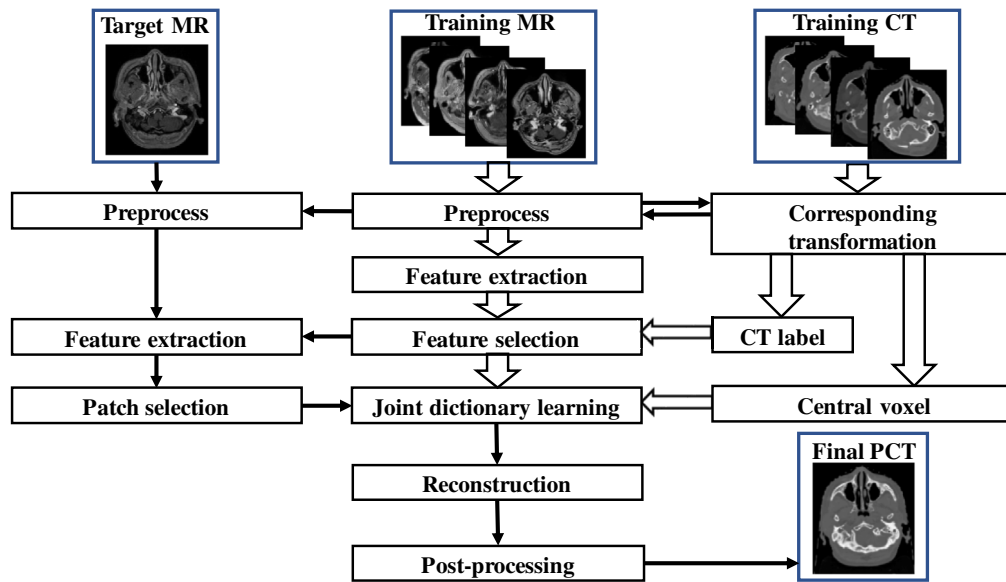


Fig. 1 The brief architecture of the proposed method.

and multivariate image modeling in transverse plane are shown in Fig. 2, where each feature map's voxel is constructed by central feature of feature vector of 3-D patch centered at that voxel's position.

2.4 Feature Selection

We denote $F = \{f(x_1), f(x_2), \dots, f(x_n)\}$ as the extracted features of MR image patches centered at each position i , $X = \{x_1, x_2, \dots, x_n\}$ is the set of CT voxels at each position

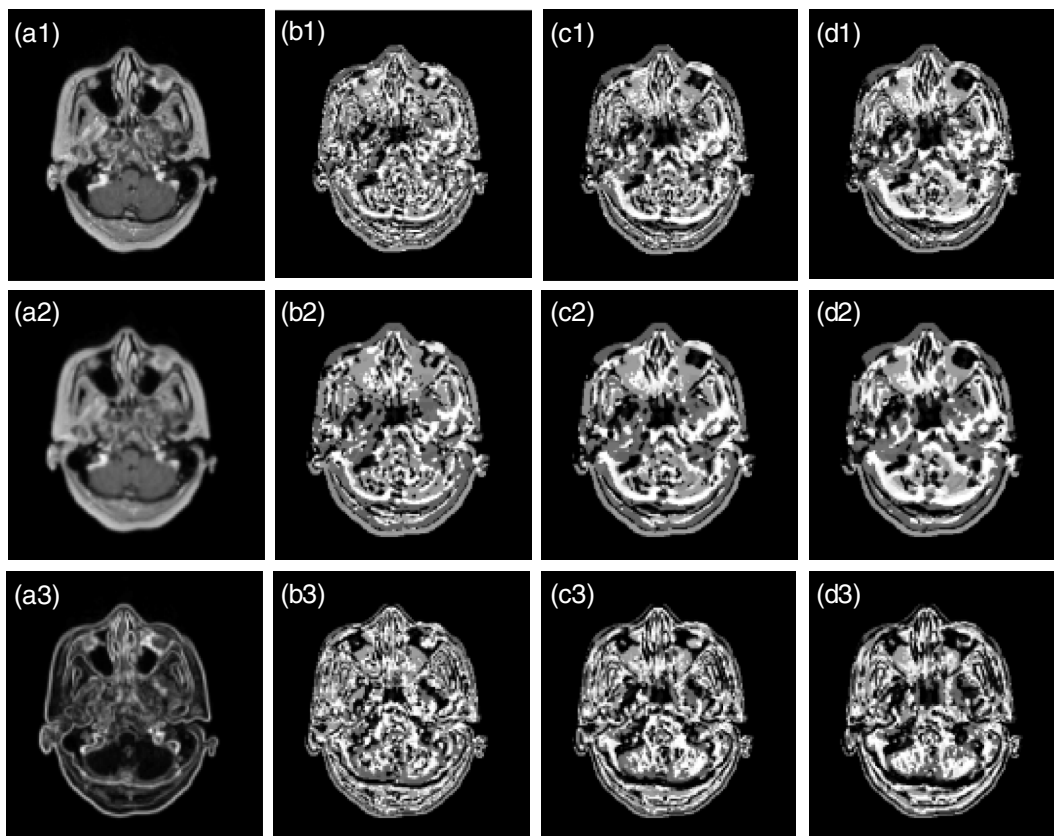


Fig. 2 Example of 3-D RILBP feature maps. The (a1–a3) are the image sets including original MRI image, mean of MRI and standard division of MRI. The corresponding feature maps are in the (b1–b3), (c1–c3), and (d1–d3). In (b1–b3), a sphere with radius 3 mm and cubes of size $3 \times 3 \times 3 \text{ mm}^3$ were used. In (c1–c3), the sphere radius was 5 mm and cube size was $5 \times 5 \times 5 \text{ mm}^3$. In (d1–d3), the sphere radius was 7 mm and cube size was $7 \times 7 \times 7 \text{ mm}^3$.

i. Recent studies^{50–52} have shown the potential drawback of using extremely high-dimensional features for existing learning methods. One major concern is that high-dimensional training data may contain bias and redundant components, which could over fit the learning models and in turn affect the final learning robustness. To address this challenge, we propose a feature selection method using a logistic regression function, as feature selection can be regarded as a binary regression task with respect to each dimension of the original feature.^{50,51} Further, the goal of feature selection is to identify and group a small subset of the most informative features. Thus, our feature selection is accomplished by implementing the sparsity constraint in the logistic regression, i.e., minimizing the following LASSO energy loss function:

$$\sum_{j=1}^n \log(1 + \exp\{-l[w^T f(x_j) + b]\}) + \mu \sum_{i=1}^m |w_i|/\beta_i, \quad (1)$$

where $f(x)$ denotes the original feature, w is a sparse coefficient vector that denotes a binary coefficient, with nonzero denoting that the corresponding features are relevant to the anatomic classification, and zero denoting that the nonrelevant features are eliminated during the classifier learning process. b denotes the intercept scalar. $l(\cdot)$ is an anatomical binary labeling function. β_i is a optimization scalar and μ is a regularization parameter. n denotes the number of voxel samples x_j and m denotes the length of vector w . β is computed by discriminative power, i.e., Fisher's score⁴⁷ for each feature of feature vector.

The optimization tries to punish the features that have smaller discriminative power, i.e., uninformative features are eliminated from training the learning-based model. To implement this punishment, β_i is used to be divided by $|w_i|$. The optimal solution of intercept scalar b and w_i can be estimated by Nesterov's⁵³ method. Denoted as $f_s(x)$, the informative features corresponding to the nonzero entries in w were selected, which have superior discriminatory power in distinguishing bone, air, and

soft-tissues from each other. We defined $f_s(x)$ as the MRI patch's anatomic signature.

Two-step feature selection was applied by first labeling air and nonair materials in first round and then labeling bone and soft-tissue in the second round. The air, soft-tissue, and bone labels, as shown in Fig. 3, are roughly segmented by fuzzy C-means clustering method. Structures adjacent to both bone and air are difficult to accurately disambiguate in comparison with voxels of other regions. Therefore, more samples were drawn from the bone and air regions during the feature selection process. To balance the three regional samples, ratio parameters were used to control the ratio of the number of samples drawn from air, bone, and soft-tissue regions.

An example is given in Fig. 4, in which (a) and (b) show training paired MRI and CT image with two types of samples. The samples of the bone region are indicated by red asterisk whereas those selected from soft-tissue regions are indicated by green circles. Figure 4(c) shows a scattered plot of the samples corresponding to the randomly selected two features without feature selection, whereas Fig. 4(d) shows that of the samples corresponding to the two top-ranked features, which are evaluated by Fisher's score,⁴⁷ after feature selection. It can be seen with feature selection that bone regions can be distinguished from the soft tissues.

2.5 Pseudo Computed Tomography Synthesis

After preselecting the most informative feature as an anatomic signature, we propose a joint dictionary learning framework to handle the challenge of large-scale data training. The aim was to estimate the target CT intensity from a MRI in a patch, or a group of voxels, wise local self-similarity fashion using the joint dictionary learning and sparse representation technique. If a patch (with a size of $5 \times 5 \times 5$ voxels) in the new arrival MR image is similar to a certain patch of training MR images, then the corresponding CT central voxel for these two MR

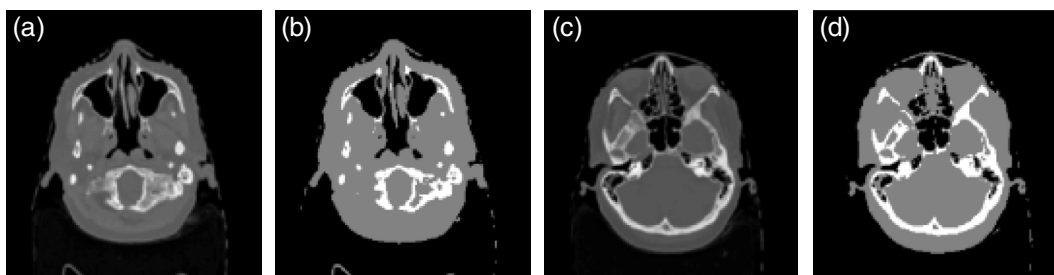


Fig. 3 Example of the CT image label. (a) and (c) The different slices of CT image, (b) and (d) the corresponding labels, where black area denotes the air label, the white area is the bone label and the gray area is the soft-tissue label.

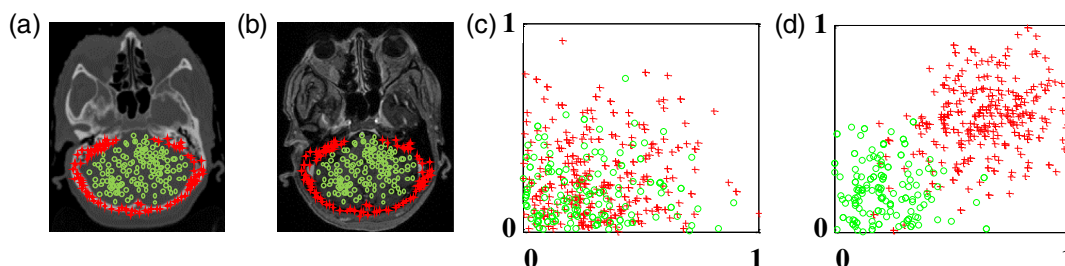


Fig. 4 An example of feature selection.

patches were highly correlated. Each patch of the new arrival MRI image could identify several similar instances from the properly selected training MR patches via a sparse learning technique when the training data set is sufficient enough. Therefore, in this paper the correlation of MR anatomic signature and CT central voxels can be constructed by joint dictionary learning, which tracks the same sparse coefficients of both MR patches and CT intensities. To this end, the framework involves the training and predicting stages. In the training stage, the following two steps are used to efficiently learn the correlation between MR anatomic signature and CT central voxel within the selected similar templates.

1. We selected a small number of similar local training MR patches and their corresponding CT central voxels in distinctive regions as the dataset for our next training step.
2. For the small set of training data, the coupled dictionaries were adaptively trained to sparsely represent the MR anatomic signature and CT intensity by joint dictionary learning, where each dictionary column is initialized by MR signature and their corresponding CT intensities.

The prediction stage consists of the following two steps:

1. For each target MR patch, we extracted similar training MR patches and the corresponding coupled dictionary so that the target patch can be sparsely represented by the selected dictionary.
2. The target CT voxel was predicted by applying sparse reconstruction of training CT intensities dictionary (i.e., a linear combination of CT dictionary columns) using the previously obtained sparse coefficients.

2.5.1 Patch selection based on local similarity

Before constructing the coupled dictionary, patch selection is needed among all candidate MR patches to reduce the computational complexity and to enhance the prediction accuracy by excluding the irrelevant patches based on reasonable similar metrics. Specifically, the criterion for patch similarity is defined between two MRI patches I_i^{MR} and I_j^{MR} by a reflecting weight w_{ij} , which can be computed by calculating distance between both CT intensity and MRI anatomic signature

$$w_{ij} = \exp[-(\|I_i^{\text{MR}} - I_j^{\text{MR}}\|_2^2 + \rho\|f_i^{\text{MR}} - f_j^{\text{MR}}\|_2^2)/2S\beta\delta], \quad (2)$$

where S is the number of voxels in the target patch, β is a smoothing parameter, δ is the standard division of all the patches and features, and ρ denotes the balance parameter of the similarity between anatomic signature and image intensity. First, the new arrival MRI is normalized by mean and standard deviation. Then, the training set D_i of the new MRI patch's central position i can be obtained by sorting the K -nearest patches within a specific searching region $N_{k_x \times k_y \times k_z}(i)$

$$D_i = \{(f_j^{\text{MR}}, y_j^{\text{MR}}) | j \in \text{sort}(w_{ij}, \text{descend}) \cap N_{k_x \times k_y \times k_z}(i)\}, \quad (3)$$

where w_{ij} reflecting the local similarities between new MRI patch I_i^{MR} and training MRI patch I_j^{MR} . w_{ij} is utilized in our next step of joint dictionary learning.

2.5.2 Joint dictionary learning

Let $\{(f_i^{\text{MR}}, I_i^{\text{MR}}, y_i^{\text{CT}}) | i = 1, 2, \dots, n\}$ denotes a database containing a set of anatomic signature, MRI patches, and corresponding CT central voxels: where $f_i^{\text{MR}} \in R^m$ denotes the anatomic signature, R^m denotes the multidimensional space, I_i^{MR} and $y_i^{\text{CT}} \in R$ denotes the corresponding MRI patch and CT central voxel intensity value, respectively.

To train the locally-adaptive coupled dictionary for each preselected training database, we initialized the first iteration of the dictionary with each coupled column comprising a pair of similar MRI anatomical signature and a corresponding central voxel of CT. Let $F = \{\omega \cdot f_i^{\text{MR}} | i = 1, 2, \dots, n\}$ denotes the set of weighted anatomic signatures of training MRI patches and $Y = \{y_i^{\text{CT}} | i = 1, 2, \dots, n\}$ denotes the set of corresponding CT intensities. The element-wise product of ω is obtained by Fisher's score⁴⁷ as shown in Sec. 2.4. The goal of applying omega is to further enforce the influence of the most discriminative signatures by adaptive weighting parameter. Then, the initialization of coupled dictionary is generated as $(D^{\text{MR}}, D^{\text{CT}}) = (F, Y)$.

In the following iterations, supposing we have the roughly constructed dictionary $D^{\text{MR}} \in R^{m \times n}$ of MRI signatures and $D^{\text{CT}} \in R^{1 \times n}$ of CT intensities on training data from the previous iteration, we utilize the joint dictionary learning method to simultaneously update D^{MR} and D^{CT} with the core idea that the sparse representation tracks the same sparse coefficients of both MR signatures and CT intensities. Let $W = \{w_i | i = 1, 2, \dots, n\}$ denotes the weights corresponding to each MRI patch I_i^{MR} computed by local self-similarity in Eq. (2), where $w_i = \{w_{ij} | j \in N_{k_x \times k_y \times k_z}(i) \setminus \{i\}\}$. Then, each training MRI anatomic signature and its corresponding CT intensity can be sparsely represented by a pair of coupled D^{MR} and D^{CT} , respectively. Let $A = \{\alpha_i | i = 1, 2, \dots, n\}$ denotes the set of sparse coefficients. Then the joint dictionary learning framework can be described as follows:

$$(D^{\text{MR}}, D^{\text{CT}}, A) = \arg \min_{D^{\text{MR}}, D^{\text{CT}}, A} \|F - D^{\text{MR}}A\|_F^2 + \mu \|Y - D^{\text{CT}}A\|_2^2 + \tau \|WA\|_{2,1}, \quad (4)$$

where τ is the regularization parameter of column sparse term, μ denotes the parameter of penalization, $\|A\|_{2,1} = \sum_{i=1}^n \|A^i\|_2$ is the joint $l_{2,1}$ -norm, and A^i is i 'th row of the matrix A . Equation (4) can be solved by ADM method.⁴⁵

2.5.3 Prediction

For each patch I_{new} of a new arrival MRI patch, the sparse representation α_{new} of a new arrival MRI patch anatomic signature under dictionary D^{MR} can be calculated by a BP optimization⁴⁶

$$\alpha_{\text{new}} = \arg \min_x \|f_{\text{new}}^{\text{MR}} - D^{\text{MR}}\alpha\|_2^2 + \tau \|\alpha\|_1, \quad (5)$$

where $f_{\text{new}}^{\text{MR}}$ denotes the anatomic signature of each new MRI patch. Finally, with the corresponding dictionary of CT D^{CT} , we utilized the coefficient α_{new} to reconstruct the

PCT intensity by a linear combination of dictionary D^{CT} columns

$$y_{\text{new}}^{\text{PCT}} = D^{\text{CT}} \alpha_{\text{new}}. \quad (6)$$

Our proposed algorithm is summarized in Algorithm 1. The default parameter setting of Algorithm 1 is shown in Table 1.

Algorithm 1 Pseudo CT prediction using anatomic signature and joint dictionary learning.

1. For each training MR image patch I_i^{MR} , extract the multiscale and multilevel 3-D RILBP feature f_i^{MR} .
2. Use fuzzy C-means to segment each corresponding CT I_i^{CT} into three labels that represent the CT values in the range of bone, soft-tissue and air.
3. For the coupled training MRI features and corresponding CT labels, apply LASSO logistic regression Eq. (1) to select the informative features as an anatomic signature.
4. Use Fisher's score⁴⁷ to measure the discriminant power of the anatomic signature to separate bone, air and soft-tissue, and normalize to ω such that $\sum \omega_i = 1$.
5. For each new arrival MRI image patch I_i^{MR} , select its similar patches I_j^{MR} from the training patches by computing the weight w_{ij} from Eq. (2), and sorting the K -nearest patches by Eq. (3).
6. For the anatomic features f_j^{MR} of I_j^{MR} and corresponding CT voxels y_j^{CT} , initialize the coupled dictionaries: $\hat{D}^{\text{MR}} = \{f_{ij}^{\text{MR}} | j \in N_{k_x \times k_y \times k_z}(i) \setminus \{i\}\}$ and $\hat{D}^{\text{CT}} = \{y_{ij}^{\text{CT}} | j \in N_{k_x \times k_y \times k_z}(i) \setminus \{i\}\}$.
7. Use joint dictionary learning [Eq. (4)] to train the coupled dictionary $(D^{\text{MR}}, D^{\text{CT}})$.
8. For each new arrival MRI patch, use Eq. (5) to obtain the sparse representation of learned dictionary D^{MR} , and then use Eq. (6) to reconstruct the pseudo CT intensity.

Table 1 Default parameter setting.

Parameter	Default value	Meaning
Window size	$9 \times 9 \times 9 \text{ mm}^3$	MRI searching region size
Patch size	$5 \times 5 \times 5 \text{ mm}^3$	MRI patch size (for feature extraction)
n	25	Number of similar patches
m	63	Length of anatomic signature
μ	0.05	Regularization parameter in Eq. (1)
β	1	Smoothing parameter of weight estimation in Eq. (2)
ρ	0.5	Balancing parameter in Eq. (2)
μ	1	Penalty parameter of joint dictionary learning in Eq. (4)
τ	0.5	Regularization parameter of joint dictionary learning in Eq. (4)

3 Experiments

3.1 Datasets and Quantitative Measurements

To test our prediction method, we applied the proposed method to 14 paired brain MR and CT images. All patients' MRI ($1.0 \times 1.0 \times 1.0 \text{ mm}^3$) and CT ($1.0 \times 1.0 \times 1.0 \text{ mm}^3$) data were acquired using a Siemens MR and CT scanner. To quantitatively characterize the prediction accuracy, we used three widely used metrics: mean absolute error (MAE), peak signal-to-noise ratio (PSNR), normalized cross correlation (NCC), and similarity index (SI)⁵⁴ for air, soft-tissue and bone regions. MAE is used to measure how close forecasts or predictions are to the actual outcomes. PSNR is an engineering term for the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. NCC is a measure of similarity between two series as a function of the displacement of one relative to the other. SI has been used to evaluate the image quality of nonrigid image registration. The three metrics are defined as

$$\text{MAE} = |I_{\text{CT}} - I_{\text{PCT}}|/C, \quad (7)$$

$$\text{PSNR} = 10 \log_{10} \left(\frac{Q^2}{\|I_{\text{CT}} - I_{\text{PCT}}\|_2^2 / C} \right), \quad (8)$$

$$\text{NCC} = \frac{1}{C} \sum_{x,y,z} \frac{[I_{\text{CT}}(x,y,z) - \mu_{\text{CT}}][I_{\text{PCT}}(x,y,z) - \mu_{\text{PCT}}]}{\sigma_{\text{CT}} \sigma_{\text{PCT}}}, \quad (9)$$

$$\text{SI} = 2 \times \frac{n\{A_{\text{CT}} \cap A_{\text{PCT}}\}}{n\{A_{\text{CT}}\} + n\{A_{\text{PCT}}\}}, \quad (10)$$

where I_{CT} is the ground-truth CT image, I_{PCT} is the corresponding PCT image, Q is the maximal intensity value of I_{CT} and I_{PCT} , and C is the number of voxels in the image. μ_{CT} and μ_{PCT} are the mean of CT and PCT image, respectively. σ_{CT} and σ_{PCT} are the standard deviation of CT and PCT image, respectively. A_{CT} and A_{PCT} denote the binary mask (air, soft-tissue, and bone) of CT and PCT image, respectively. We set intensity value < -400 HU as air, > 300 HU as bone, and the rest as soft-tissue. $n\{A\}$ is the number of elements in set A . In general, a better prediction has lower MAE, and higher PSNR, NCC, and SI values.

3.2 Parameter Setting

To compare the influence of multiple parameters within our proposed algorithm (shown in Table 1) we evaluated how variation of one parameter affects performance while fixing the other parameters to their default settings in Table 1. Our parameter testing experiment was performed by testing one patient and training the other 13 patients. This was done by comparing the three metrics in Eqs. (7)–(9). We tested the regularization parameter μ of Eq. (1) over the range of 0.01 to 0.1 and fixed the other parameter as default settings in Table 1. The performance was not sensitive around $\mu = 0.05$. The smoothing parameter β and balancing parameter ρ in Eq. (2) were empirically set to 1 and 0.5, respectively. These parameters were decided by their performance on our experiments. In Eq. (4), the parameters μ and τ were initialized by 1 and 0.5, in the

Table 2 The performance with different patch size and window size.

Patch size	Window size		
	5 × 5 × 5	7 × 7 × 7	9 × 9 × 9
3 × 3 × 3	PSNR = 22.7 dB	PSNR = 22.9dB	PSNR = 22.7 dB
	NCC = 0.93	NCC = 0.93	NCC = 0.92
	MAE = 76.1 HU	MAE = 79.3 HU	MAE = 83.6 HU
5 × 5 × 5	N/A	PSNR = 23.7 dB	PSNR = 23.7 dB
		NCC = 0.93	NCC = 0.93
		MAE = 76.1 HU	MAE = 75.9 HU
7 × 7 × 7	N/A	N/A	PSNR = 23.1 dB
			NCC = 0.93
			MAE = 76.9 HU

following iterations, these two parameters were updated by ADM as previously recommended.⁴⁵

We also fixed the patch size and tested different window sizes, and vice versa. As seen in Table 2, the best MAE and NCC were obtained with a patch size of 5 × 5 × 5 mm³ and a window size of 9 × 9 × 9 mm³ (Table 2).

Figures 5(a1)–5(a3) show PSNR, MAE, and NCC metrics as a function of number of similar patches, whereas the other parameters are fixed as in default setting in Table 1. From these curves, 23 similar patches are sufficient for a good prediction. However, since MAE was the highest priority metric that

was optimized, the absolute change in PSNR from 23 to 25 similar patches holds less weight than MAE, which resulted in 25 similar patches being the most optimal number for prediction.

The number of selected features also influences the learning performance. If most features are eliminated by feature selection, the precision of PCT prediction may be diminished due to insufficient information. Conversely, if the majority of features are selected, uninformative data may also affect the accuracy of the proposed method. To measure the performance as determined by the number of selected features, we evaluated the PSNR, NCC, and MAE metrics as shown in Figs. 5(b1)–5(b3), and fixed the other parameter as shown in Table 1. One can observe that the optimal number of selected features is 63.

3.3 Contribution of Feature Selection and Joint Dictionary Learning

In Sec. 2, we introduced two enhancements over the classic feature-based method, i.e., feature selection and joint dictionary learning. Figure 6 shows detailed visual comparisons between classic feature-based (CF), feature selection-based (FS), and feature selection with jointly dictionary learning-based (FSJDL) methods. Table 3 quantifies the above observations using 14 pairs of MR-CT leave-one-out experiments. We can visually observe the result in Fig. 6 that the FSJDL method best preserves continuity, coalition, and smoothness in the prediction results. As prediction results between methods primarily differ in regions were bone is adjacent to air, visual results within such regions are shown in Figs. 6(b2)–6(d2) and 6(b5)–6(d5). Table 4 shows the consistent improvement in FS over the CF method, and the modest improvement of FSJDL over the FS and CF methods. The prediction estimated by FS is statistically superior to the CF method for MAE, NCC, and PSNR (*t*-test, *p* < 0.001).

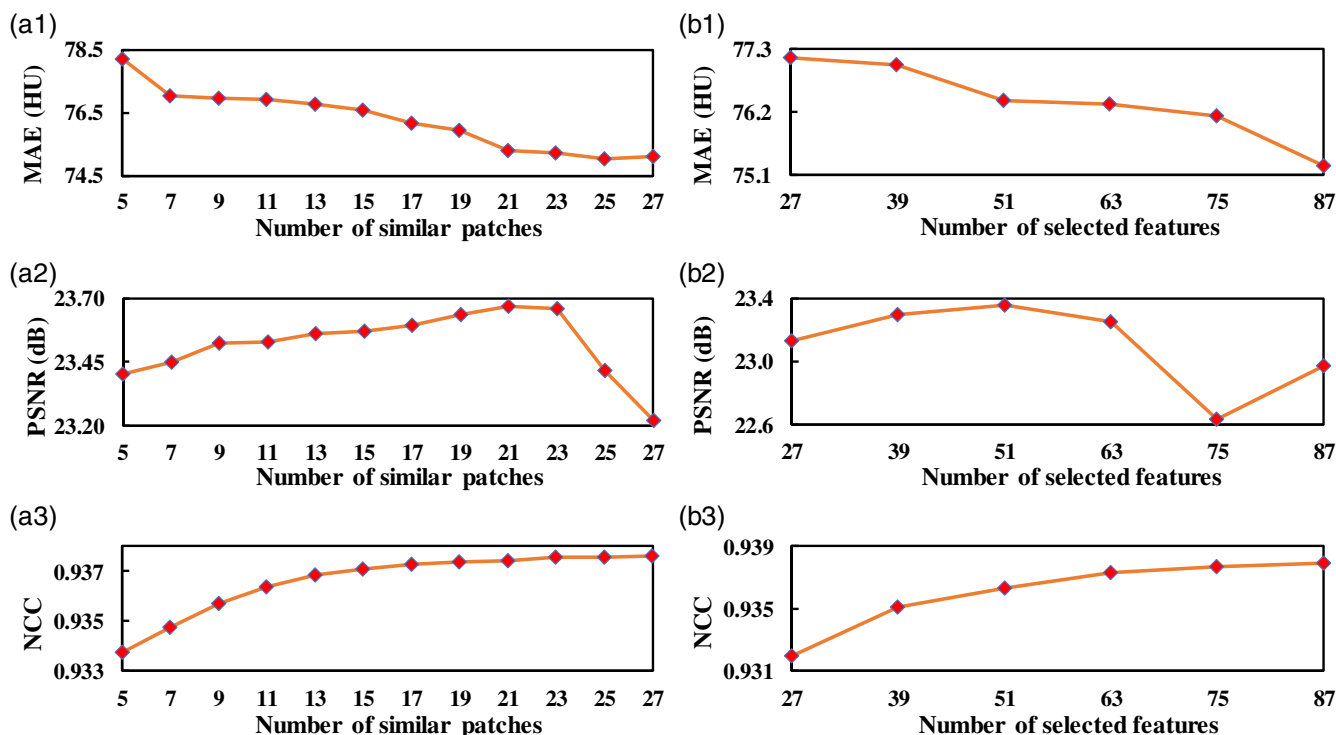


Fig. 5 (a1–a3) The MAE, PSNR, and NCC performance with different number of similar patches and (b1–b3) the MAE, PSNR, and NCC performance with different number of selected features.

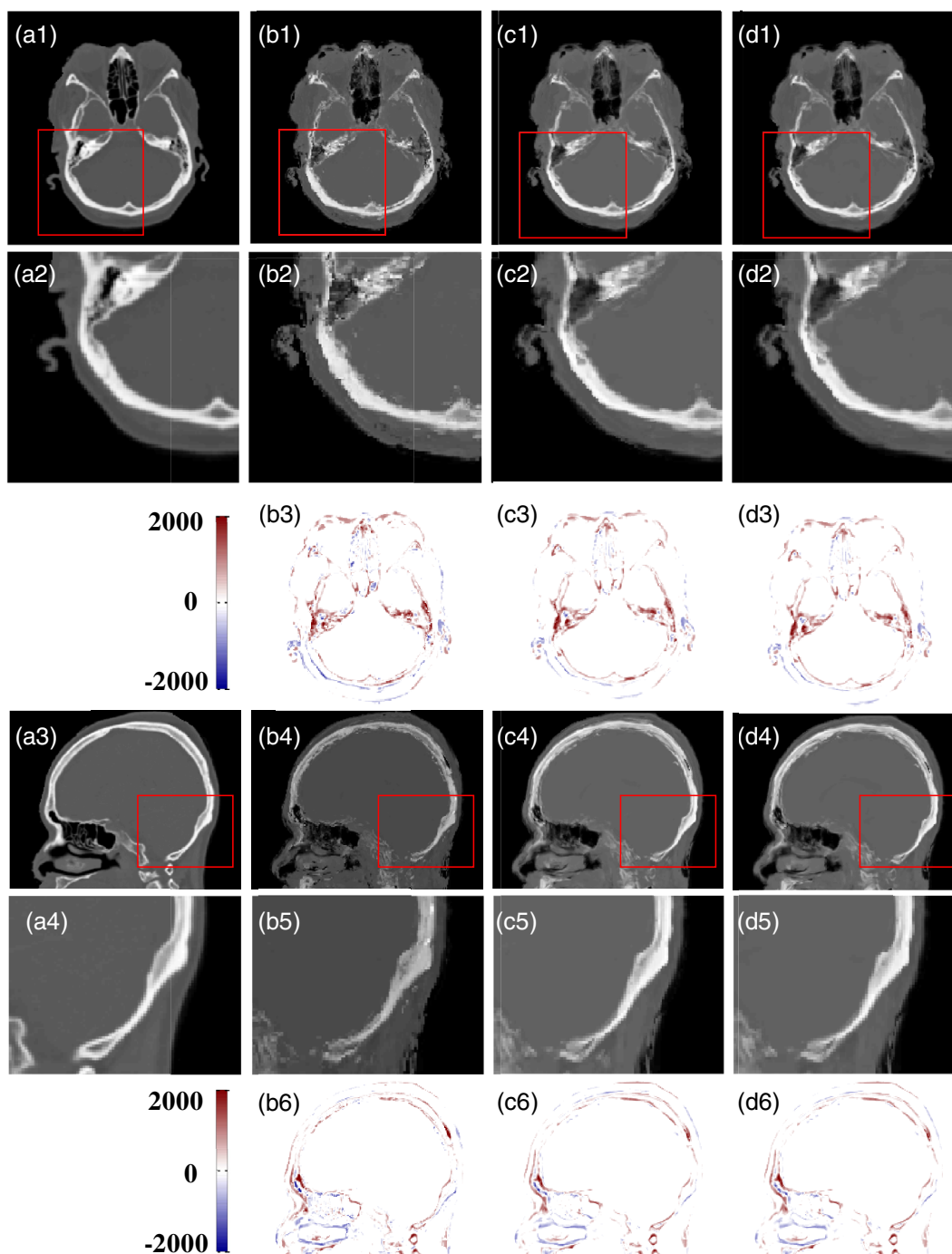


Fig. 6 Comparison of the PCT estimated by CF, FS, and FSJDL, (a1) and (a3) ground truth CT, (b1–d1) and (b4–d4) are the PCT result generated by CF, FS, and FSJDL, respectively. (a2–d2) The zoomed regions indicated by solid red boxes in (a1–d1); (a4–d4) are the zoomed regions indicated by solid red boxes in (a3–d3). (b3–d3) are difference image between ground truth (a1) and the PCT estimated by CF, FS, and FSJDL. (b6–d6) Difference image between ground truth (a3) and the PCT generated by CF, FS, and FSJDL.

Table 3 Numerical results with compared methods.

Method	MAE (HU)	PSNR (dB)	NCC
CF	102.8 ± 22.7	21.6 ± 1.2	0.87 ± 0.04
FS	96.4 ± 23.8	22.0 ± 1.4	0.88 ± 0.04
FSJDL	82.6 ± 26.1	22.4 ± 1.9	0.91 ± 0.04

Also, the prediction estimated by FSJDL is significantly superior to the FS method for MAE and NCC ($p < 0.001$).

3.4 Comparison of State-of-the-Art Dictionary Learning-Based Methods

We compared our method with two state-of-the-art dictionary-based approaches: an intensity-based method (Int)⁴² and a fast patch-based method (FP).⁴³ The parameters used in these two

Table 4 Numerical results of the proposed and state-of-the-art dictionary learning methods.

Method	MAE (HU)	PSNR (dB)	NCC	SI (Air)	SI (Soft-tissue)	SI (Bone)
Int	148.9 ± 42.3	18.3 ± 0.9	0.80 ± 0.05	0.80 ± 0.09	0.57 ± 0.12	0.51 ± 0.13
FP	125.5 ± 24.5	19.1 ± 1.1	0.80 ± 0.05	0.93 ± 0.02	0.72 ± 0.07	0.56 ± 0.18
FSJDL	82.6 ± 26.1	22.4 ± 1.9	0.91 ± 0.04	0.98 ± 0.01	0.88 ± 0.03	0.69 ± 0.08
FSJDL versus Int (<i>p</i> -value)	<0.001	<0.001	<0.001	<0.001	<0.001	0.002
FSJDL versus FP (<i>p</i> -value)	<0.001	<0.001	<0.001	<0.001	<0.001	0.036

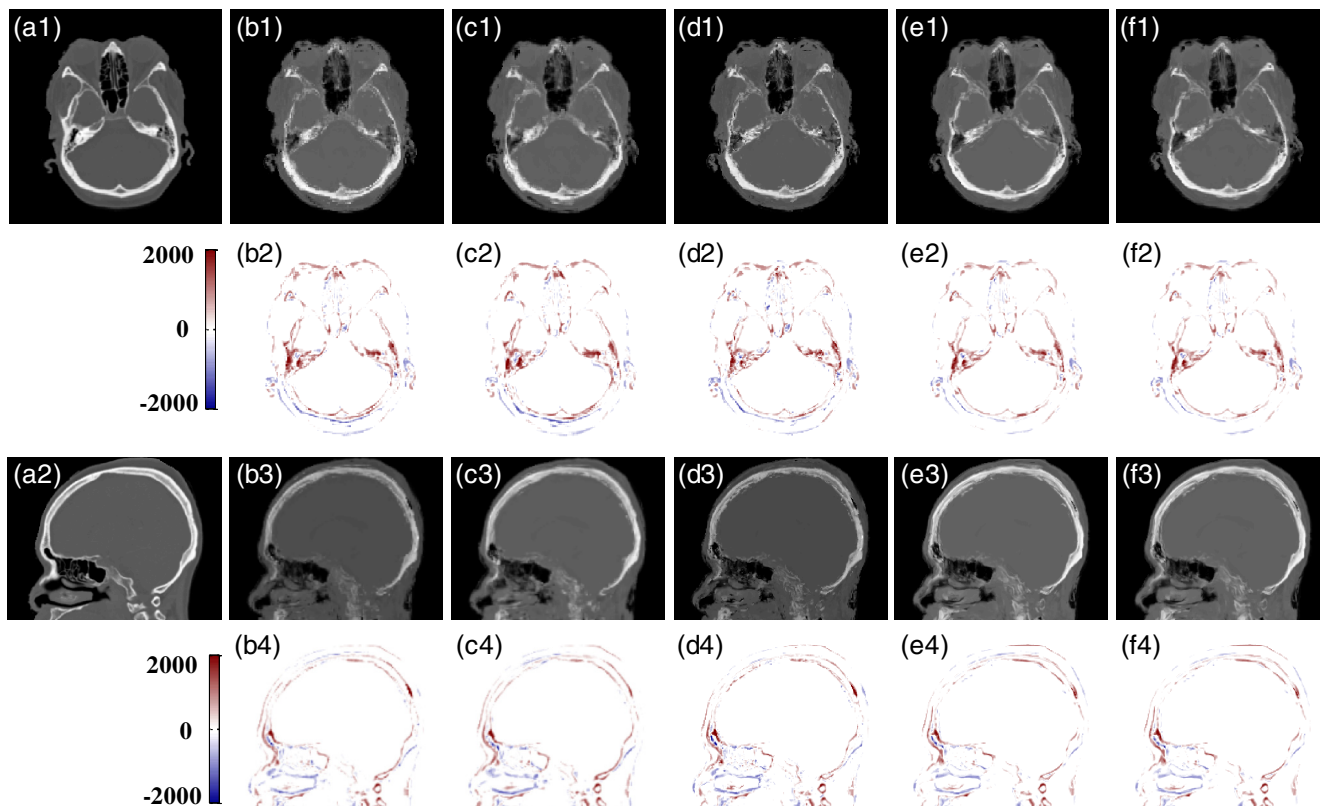


Fig. 7 Comparison with two state-of-the-art methods: (a1–a2) ground truth, (b1–b4) intensity-based method (Int), (c1–c4) fast patch-based method (FP), (d1–d4) classic feature-based method (CF), (e1–e4) feature selection-based method (FS), and (f1–f4) feature selection with joint dictionary learning-based method (FSJDL).

methods were set based on their best performance. The prediction results are shown in Fig. 7. The CT images generated by our method are similar to ground truth. Our method outperforms these two methods, specifically with higher similarity in shape and lower difference values. The Int method often results in blurred predictions due to simple averaging of the similar patches. Although the FP method works better than the Int method, its prediction remains noisy. Table 4 quantifies the superiority of our method over the other two in terms of MAE, PSNR, NCC, and SI. Our method’s average MAE of 82.6 HU is significantly better than the 148.9 and 125.5 HU generated by the Int and the FP methods, respectively. Figure 8 shows the detailed prediction performance of the comparison method through histogram analyses of the entire leave-one-out template. It can be seen that for some patients, overall

performance is inferior to that of others, presumably owing to lack of database diversity.

4 Discussion

In our study, we have presented a dictionary-learning-based method of estimating a PCT image from a standard MR image. We introduced an anatomic signature based on feature selection into a dictionary-learning-based method to improve the prediction accuracy of the resultant PCT. Local self-similarity was incorporated to address the challenge of large-scale training data training. Joint dictionary learning was applied to learn the coupled dictionary of MRI anatomic signatures and CT intensities. We evaluated the clinical application of our method on MR-CT brain datasets and compared its performance

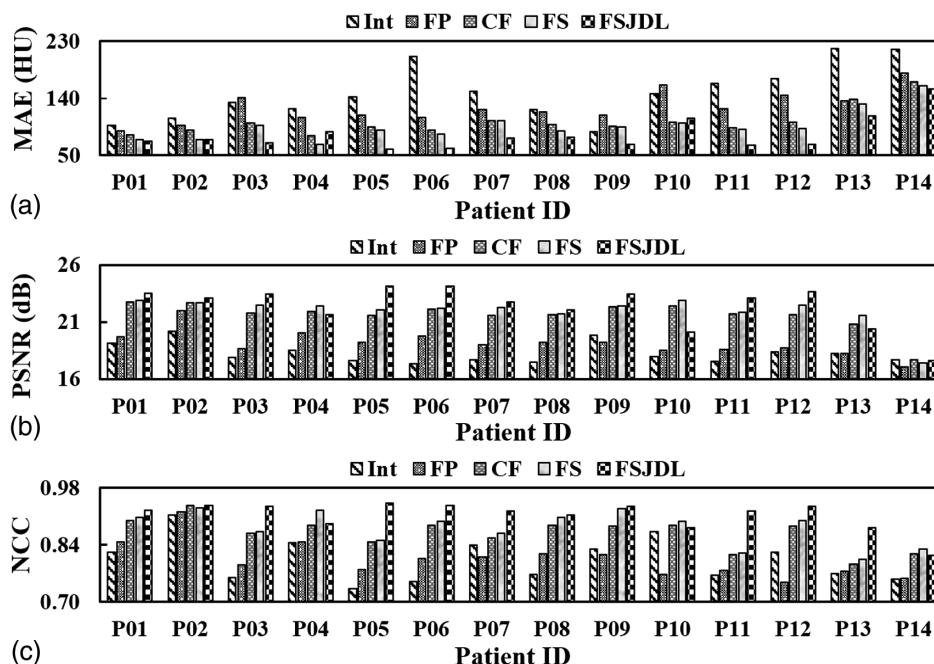


Fig. 8 (a–c) The MAE, PSNR, and NCC of different methods for each enrolled patient, respectively.

with the true CT image (ground truth) and two state-of-the-art dictionary-learning-based methods.

The efficacy of using 3-D RILBP was visually demonstrated in Figs. 7(d1)–7(d4) when compared with [Figs. 7(c1)–7(c4)] and numerically demonstrated by detailed metrics comparing FP and CF methods in Fig. 7. We observed that 3-D RILBP features notably improved the prediction accuracy of the PCT. Figure 2 shows that multiradii level sensitivity and multi-derivate image modeling provided features containing local texture, edges and spatial information derived from a 3-D MR image patch. Thus, we concluded that 3-D RILBP features can balance image deformation invariance and sensitivity to small structural changes.

The efficacy of using an anatomic signature was demonstrated via the comparison between Figs. 6(c1)–6(c6), compared with Figs. 6(b1)–6(b6), Figs. 7(e1)–7(e4) versus Figs. 7(d1)–7(d4) and Table 4. The PCT generated from FP method was superior compared with that generated with the CF method. Our results showed that the anatomic signature substantially improved the prediction accuracy, both qualitatively and quantitatively. It can be concluded that the use of anatomic signature identified by feature selection can diminish the influence of bias and redundant information contained by high-dimensional training data and preserve the informative and highly relevant features for prediction performance.

Figures 6(d1)–6(d6) compared with Figs. 6(c1)–6(c6), and Figs. 7(f1)–7(f4) compared with Figs. 7(e1)–7(e4), and Fig. 8 FSJDL compared with the FS method validated the efficacy of the local self-similarity and joint dictionary learning framework. The FSJDL method was a significant improvement to FS method. Since the training was based on a small set of similar patches, a CT image could be predicted in about 2 h on 3.00 GHz Intel Xeon 8-core processor, compared to about 1 week required for competing methods in the same experimental environment. Thus, joint dictionary learning drastically reduces computation complexity.

Our method was substantially better than two state-of-the-art dictionary learning methods as evaluated by the local texture, edge and spatial information anatomical signature extraction, the discriminative feature selection, and the joint dictionary learning framework. As shown in Fig. 7, our system could predict PCT images that are similar to ground-truth values and could effectively capture minute changes in electron densities. As shown in Table 4 and Fig. 8, our proposed method significantly improved the performance of PCT. The MAE between our generated PCT and ground truth CT is <85 Hounsfield unit (HU). The previous work¹⁴ has shown that an HU error on the order of ~100 did not affect the dosimetric accuracy of intensity modulated radiation treatment planning based on PCT. This shows that PCT generated from MRI is a valid clinical application that can be applied to the current workflow of patient treatment planning.

Although we have proposed an accurate method for estimating PCTs derived from brain T1-weighted MR, this work still has several limitations. First, we did not apply this method to other body sites, such as abdomen and lung. For nonrigid structures, we will have to use a deformable MRI-CT registration to deal with the intrasubject mismatch between MRI and CT. Second, a dosimetric comparison between plans generated from the simulation CT and its corresponding predicted CT has not yet been explored. Third, we did not compare our algorithm with the other state-of-the-art machine-learning-based methods. The goals here are proving the efficiency of the feature selection in manufactural feature-based dictionary learning method, and improving the recent dictionary-learning-based methods for PCT estimation. Last, we tested our algorithm on 14 patients’ brain; testing in a larger population can prove the robustness of our algorithm.

5 Conclusion

Here, we proposed anatomic signature and joint dictionary learning-based methods to improve the accuracy of pseudo

CT prediction. The anatomic signature was generated by identifying the most informative and discriminative features within the 3-D RFLBP features extracted from multiradii level sensitivity and multidervate image modeling-based MR images. The feature selection was implemented by LASSO and then the Fisher's score was adopted to evaluate the relevance of such feature anatomically and statistically. To feasibly implement the specific dictionary learning for each patch of an arriving new MRI, a bounded searching region with most similar neighbors was obtained by computing the local self-similarity. Then, a joint dictionary learning method was incorporated to simultaneously train the sparse representation coupled MRI and CT dictionaries of either MRI anatomic signature or CT intensities. Afterward, the sparse coefficients of new MRI patch signature were optimized by BP methods under the well-trained MRI dictionary. Finally, we used these coefficients to reconstruct the pseudo CT intensity under the well-trained CT dictionary. Experimental results showed that our method can accurately predict pseudo CT images in various scenarios, even in the context of large shape variation, and outperforms two state-of-the-art dictionary-learning-based methods.

Disclosures

The author declares no conflicts of interest.

Acknowledgments

This research was supported in part by the National Cancer Institute of the National Institutes of Health under Award No. R01CA215718, the Department of Defense (DoD) Prostate Cancer Research Program (PCRP) Award No. W81XWH-13-1-0269 and Dunwoody Golf Club Prostate Cancer Research Award, a philanthropic award provided by the Winship Cancer Institute of Emory University.

References

1. J. M. Edmund and T. Nyholm, "A review of substitute CT generation for MRI-only radiation therapy," *Radiat. Oncol.* **12**, 28 (2017).
2. S. Devic, "MRI simulation for radiotherapy treatment planning," *Med. Phys.* **39**(11), 6701–6711 (2012).
3. D. Andreasen et al., "Patch-based generation of a pseudo CT from conventional MRI sequences for MRI-only radiotherapy of the brain," *Med. Phys.* **42**(4), 1596–1605 (2015).
4. M. Kapanen and M. Tenhunen, "T1/T2*-weighted MRI provides clinically relevant pseudo-CT density data for the pelvic bones in MRI-only based radiotherapy treatment planning," *Acta Oncol.* **52**(3), 612–618 (2013).
5. J. Uha et al., "MRI-based treatment planning with pseudo CT generated through atlas registration," *Med. Phys.* **41**(5), 051711 (2014).
6. J. Sjolund et al., "Generating patient specific pseudo-CT of the head from MR using atlas-based regression," *Phys. Med. Biol.* **60**(2), 825–839 (2015).
7. E. Schreiber et al., "MR-based attenuation correction for hybrid PET-MR brain imaging systems using deformable image registration," *Med. Phys.* **37**(5), 2101–2109 (2010).
8. J. A. Dowling et al., "An atlas-based electron density mapping method for magnetic resonance imaging (MRI)-alone treatment planning and adaptive MRI-based prostate radiation therapy," *Int. J. Radiat. Oncol.* **83**(1), E5–E11 (2012).
9. J. Sjolund et al., "Skull segmentation in MRI by a support vector machine combining local and global features," in *22nd Int. Conf. on Pattern Recognition*, pp. 3274–3279 (2014).
10. M. Hofmann et al., "MRI-based attenuation correction for PET/MRI: a novel approach combining pattern recognition and atlas registration," *J. Nucl. Med.* **49**(11), 1875–1883 (2008).

11. P. Khateri et al., "A novel segmentation approach for implementation of MRAC in head PET/MRI employing short-TE MRI and 2-point Dixon method in a fuzzy C-means framework," *Nucl. Instrum. Methods Phys. Res. Sect. A* **734**, 171–174 (2014).
12. P. Khateri et al., "Generation of a four-class attenuation map for MRI-based attenuation correction of PET Data in the head area using a novel combination of STE/Dixon-MRI and FCM clustering," *Mol. Imaging Biol.* **17**(6), 884–892 (2015).
13. H. Yu et al., "Toward magnetic resonance-only simulation: segmentation of bone in MR for radiation therapy verification of the head," *Int. J. Radiat. Oncol.* **89**(3), 649–657 (2014).
14. M. S. R. Gudur et al., "A unifying probabilistic Bayesian approach to derive electron density from MRI for radiation therapy treatment planning," *Phys. Med. Biol.* **59**(21), 6595–6606 (2014).
15. C. Siversson et al., "Technical note: MRI only prostate radiotherapy planning using the statistical decomposition algorithm," *Med. Phys.* **42**(10), 6090–6097 (2015).
16. X. F. Yang and B. W. Fei, "Multiscale segmentation of the skull in MR images for MRI-based attenuation correction of combined MR/PET," *J. Am. Med. Inf. Assoc.* **20**(6), 1037–1045 (2013).
17. A. P. Aitken et al., "Improved UTE-based attenuation correction for cranial PET-MR using dynamic magnetic field monitoring," *Med. Phys.* **41**(1), 012302 (2014).
18. C. Buerger et al., "Investigation of MR-based attenuation correction and motion compensation for hybrid PET/MR," *IEEE Trans. Nucl. Sci.* **59**(5), 1967–1976 (2012).
19. J. M. Edmund et al., "A voxel-based investigation for MRI-only radiotherapy of the brain using ultra short echo times," *Phys. Med. Biol.* **59**(23), 7501–7519 (2014).
20. V. Keereman et al., "MRI-based attenuation correction for PET/MRI using ultrashort echo time sequences," *J. Nucl. Med.* **51**(5), 812–818 (2010).
21. L. B. Aasheim et al., "PET/MR brain imaging: evaluation of clinical UTE-based attenuation correction," *Eur. J. Nucl. Med. Mol. Imaging* **42**(9), 1439–1446 (2015).
22. J. Cabello et al., "MR-based attenuation correction using ultrashort-echo-time pulse sequences in dementia patients," *J. Nucl. Med.* **56**(3), 423–429 (2015).
23. Y. Berker et al., "MRI-based attenuation correction for hybrid PET/MRI systems: a 4-class tissue segmentation technique using a combined ultrashort-echo-time/dixon MRI sequence," *J. Nucl. Med.* **53**(9), 796–804 (2012).
24. C. N. Ladefoged et al., "Region specific optimization of continuous linear attenuation coefficients based on UTE (RESOLUTE): application to PET/MR brain imaging," *Phys. Med. Biol.* **60**(20), 8047–8065 (2015).
25. S. H. Hsu et al., "Investigation of a method for generating synthetic CT models from MRI scans of the head and neck for radiation therapy," *Phys. Med. Biol.* **58**(23), 8419–8435 (2013).
26. C. M. Rank et al., "MRI-based treatment plan simulation and adaptation for ion radiotherapy using a classification-based approach," *Radiat. Oncol.* **8**, 51 (2013).
27. C. M. Rank et al., "MRI-based simulation of treatment plans for ion radiotherapy in the brain region," *Radiother. Oncol.* **109**(3), 414–418 (2013).
28. J. Kim et al., "Implementation of a novel algorithm for generating synthetic CT images from magnetic resonance imaging data sets for prostate cancer radiation therapy," *Int. J. Radiat. Oncol.* **91**(1), 39–47 (2015).
29. J. Korhonen et al., "Influence of MRI-based bone outline definition errors on external radiotherapy dose calculation accuracy in heterogeneous pseudo-CT images of prostate cancer patients," *Acta Oncol.* **53**(8), 1100–1106 (2014).
30. J. Korhonen et al., "A dual model HU conversion from MRI intensity values within and outside of bone segment for MRI-based radiotherapy treatment planning of prostate cancer," *Med. Phys.* **41**(1), 011704 (2014).
31. X. F. Yang and B. W. Fei, "A skull segmentation method for brain MR images based on multiscale bilateral filtering scheme," *Proc. SPIE* **7623**, 76233K (2010).
32. B. K. Navalpakkam et al., "Magnetic resonance-based attenuation correction for PET/MR hybrid imaging using continuous valued attenuation maps," *Invest. Radiol.* **48**(5), 323–332 (2013).

33. A. Johansson, M. Karlsson, and T. Nyholm, "CT substitute derived from MRI sequences with ultrashort echo time," *Med. Phys.* **38**(5), 2708–2714 (2011).
34. A. Johanson et al., "Improved quality of computed tomography substitute derived from magnetic resonance (MR) data by incorporation of spatial information—potential application for MR-only radiotherapy and attenuation correction in positron emission tomography," *Acta Oncol.* **52**(7), 1369–1373 (2013).
35. J. H. Jonsson et al., "Treatment planning of intracranial targets on MRI derived substitute CT data," *Radiother. Oncol.* **108**(1), 118–122 (2013).
36. J. H. Jonsson et al., "Accuracy of inverse treatment planning on substitute CT images derived from MR data for brain lesions," *Radiat. Oncol.* **10**, 13 (2015).
37. R. J. Li et al., "Deep learning based imaging data completion for improved brain disease diagnosis," *Lect. Notes Comput. Sci.* **8675**, 305–312 (2014).
38. A. Jog, A. Carass, and J. L. Prince, "Improving magnetic resonance resolution with supervised learning," in *IEEE 11th Int. Symp. on Biomedical Imaging (ISBI)*, pp. 987–990 (2014).
39. D. Andreasen et al., "Computed tomography synthesis from magnetic resonance images in the pelvis using multiple random forests and auto-context features," *Proc. SPIE* **9784**, 978417 (2016).
40. S. L. S. Chan et al., "Automated classification of bone and air volumes for hybrid PET-MRI brain imaging," in *Int. Conf. on Digital Image Computing: Techniques and Applications (Dicta)*, pp. 1–8 (2013).
41. T. Huynh et al., "Estimating CT image from MRI data using structured random forest and auto-context model," *IEEE Trans. Med. Imaging* **35**(1), 174–183 (2016).
42. D. Andreasen, K. Van Leemput, and J. M. Edmund, "A patch-based pseudo-CT approach for MRI-only radiotherapy in the pelvis," *Med. Phys.* **43**(8), 4742–4752 (2016).
43. A. Torrado-Carvajal et al., "Fast patch-based pseudo-CT synthesis from T1-weighted MR images for PET/MR attenuation correction in brain studies," *J. Nucl. Med.* **57**(1), 136–143 (2016).
44. S. Aouadi et al., "Sparse patch-based method applied to MRI-only radiotherapy planning," *Phys. Med.* **32**, 309 (2016).
45. S. Boyd et al., "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.* **3**(1), 1–122 (2010).
46. W. Deng, W. Yin, and Y. Zhang, "Group sparse optimization by alternating direction method," *Proc. SPIE* **8858**, 88580R (2013).
47. M. W. Ayeche and D. Ziou, "Automated feature weighting and random pixel sampling in k-means clustering for Terahertz image segmentation," in *IEEE Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW)* (2015).
48. O. Pauly et al., "Fast multiple organ detection and localization in whole-body MR Dixon sequences," *Lect. Notes Comput. Sci.* **6893**, 239–247 (2011).
49. J. Fehr and H. Burkhardt, "3D rotation invariant local binary patterns," in *19th Int. Conf. on Pattern Recognition*, Vol. 1–6, pp. 616–619 (2008).
50. M. Avalos et al., "Sparse conditional logistic regression for analyzing large-scale matched data from epidemiological studies: a simple algorithm," *BMC Bioinf.* **16**, S1 (2015).
51. S. Aseervatham et al., "A sparse version of the ridge logistic regression for large-scale text categorization," *Pattern Recognit. Lett.* **32**(2), 101–106 (2011).
52. M. Yaqub et al., "Investigation of the role of feature selection and weighted voting in random forests for 3-D volumetric segmentation," *IEEE Trans. Med. Imaging* **33**(2), 258–271 (2014).
53. Y. Nesterov, *Introductory Lectures on Convex Optimization: A Basic Course*, Springer US, New York (2003).
54. F. H. Tang and H. S. H. Ip, "Image fusion enhancement of deformable human structures using a two-stage warping-deformable strategy: a content-based image retrieval consideration," *Inf. Syst. Front.* **11**(4), 381–389 (2009).

Biographies for the authors are not available.